DOCUMENT RESUME

ED 459 696 IR 021 055

AUTHOR Miller-Whitehead, Marie

TITLE Conducting a Web Search.

INSTITUTION Tennessee Valley Educators for Excellence, Muscle Shoals, AL.
PUB DATE 2001-09-00
NOTE 10p.
AVAILABLE FROM Tennessee Valley Educators for Excellence, P.O. Box 2882, Muscle Shoals, AL 35662. Web site: http://www.tvee.org.
PUB TYPE Reports - Research (143)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Databases; Information Retrieval; Information Seeking; *Online Catalogs; *Online Searching; Search Strategies; *World Wide Web
IDENTIFIERS America Online; ERIC; Lycos; *Search Engines; Web Sites; Yahoo

ABSTRACT

Keyword and text string searches of online library catalogs often provide different results according to library and database used and depending upon how books and journals are indexed. For this reason, online databases such as ERIC often provide tutorials and recommendations for searching their site, such as how to use Boolean search strategies. Using Internet search engines to conduct research is an even more problematic task than searching online library catalogs. This paper focuses on conducting an effective search of the Internet using several popular Internet search engines. A keyword search was conducted using Excite, Yahoo, AOL. Lycos, Alta Vista and Google search engines with the same search query string, "substance abuse research," used each time. A follow-up query of ERIC was conducted using the same search terms, keeping in mind that ERIC searches yield results only for published articles in journals (including electronic journals) or published by one of the ERIC Clearinghouses. An annotated list of sites that resulted from the keyword search of "substance abuse research" by AOL, Yahoo, and Lycos is provided. The paper concludes that results of queries of search engines may be influenced by the following: content of the pages; hypertext markup language authoring skills of the Web author; design of the search engine's Web site rating method; registration of the site with search engines; selection of keywords used in the search; number of pages on the Web site and number of times the keywords appear; and intended audience and purpose of the Web site. (Author/AEF)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 459 696

CONDUCTING A WEB SEARCH

Marie Miller-Whitehead, Ph.D.

TENNESSEE VALLEY EDUCATORS FOR EXCELLENCE
TVEE.ORG

TVEE.ORG
PO Box 2882
Muscle Shoals, AL 35662

BEST COPY AVAILABLE

IR021055

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

M. Miller-Whitehead

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Conducting a Web Search

Marie Miller-Whitehead
Tennessee Valley Educators for Excellence
TVEE.ORG

Abstract

Keyword and text string searches of online library catalogs often provide different results according to library and database used and depending upon how books and journals are indexed. For this reason online databases such as ERIC often provide tutorials and recommendations for searching their site, such as how to use Boolean search strategies. Using Internet search engines to conduct research is an even more problematic task than searching online library catalogs. This paper provides some suggestions for conducting an effective search of the Internet using several popular Internet search engines.

Introduction

The proliferation of information available over the internet has proven to be both a blessing and a curse in some respects, resulting not only in information overload but also in the equally problematic selection of appropriate search engines to help find and sort out the plethora of information that is available. Those who work with web authoring and web based research know that there are now several thousand search engines in operation on the Internet. Therefore, finding and evaluating information available on the Internet has now become an independent area of research for media and information specialists, librarians, and other researchers. The researcher wishing to develop a list of resources may proceed in a variety of ways: a catalog or database of web sites may be diverse or heterogeneous, but representative of sites that the author or researcher has occasion to visit often in carrying out the duties of his or her profession, or the list may be more homogeneous of the type that a keyword search of a search engine would yield. Search engines have now progressed so that it is possible to search by quite sophisticated text strings, rather than the more limited keyword search. A counseling professional might search for links to professional organizations, to partnering clinics and organizations, to client organizations, and to local, state, and national data sources in healthcare as well as to sources of articles on healthcare in academic journals and in the popular press. Several of the links on such a list might result from a keyword search of "health care" or a search string such as "adolescent health and nutrition." If there is a particular area within healthcare that is of interest, perhaps substance abuse, once again the same principles apply. The professional will have an established network of resources, known and readily available, and will use keyword or text string searches to augment existing information by adding new links to those already established. The professional will have access to many online or electronic resources that may not be widely known or available to the

general public, including subscription database resources such as WestLaw, PsychLit and MedLine, to mention only three. With the expansion of more journals into electronic publication, many excellent peer-reviewed articles previously available only to the academic community are now being made available on the Internet to those who can find them.

Even if many resources are available only to universities and professional researchers or by subscription to a service, many others are available to the public, and there are an enormous number of Internet search engines and search engine models available for use by the researcher who wishes to find them. Each will yield somewhat different results when queried with the same keywords or search string terms. There are several reasons underlying the differences in results. A search engine may search for content of the pages, and rank order number of times the keywords appear on the page, thus web sites with more pages or with pages that use the keyword or phrases repeatedly may be ranked higher than similar sites where the specific keywords or phrases are used only once or twice. The engine may also search for what are called "metatags" that can be inserted by those who are professionals with web authoring. Thus, the results will reflect more of the web author's skill in writing of metatags than in the content of the pages. Search engines using this method can be fooled into returning results that may not be reflective of the content of the web site. Because doing research is time-consuming and time is a valuable commodity, prior planning, "doing one's homework," or the use of an intermediary media specialist can be time well spent, particularly if the search topic is either very broad or, conversely, very narrow (Brem & Boyes, 2000; Hertzberg & Rudner, 1999). Also unfortunately, the common use of acronyms in domain names and the growing number of domain extensions in use (such as .com, .net, .org, .gov., .mil, .edu) sometimes causes searches to yield many irrelevant pages, as an Internet search for "the White House" may indicate.

Methodology

Some search engines use a combination of methods to rank order query results so a comparison of results using the same keywords and different search engines was of interest in preparing an annotated list of web sites on "substance abuse research." A keyword search was conducted using Excite, Yahoo, AOL, Lycos, Alta Vista and Google search engines with the same search query string "substance abuse research" used each time. A followup query of ERIC was conducted using the same search terms, keeping in mind that ERIC searches yield results only for published articles in journals (including electronic journals) or published by one of the ERIC Clearinghouses. Few if any of the sites listed below would be found in an ERIC search, unless the organization published a book or journal listed by ERIC. While all results of the search are not reported here, the Excite search engine provided substantially

different links than those using the other web search tools. Here is an annotated list of sites that resulted from a keyword search of "substance abuse research" by AOL, Yahoo, and Lycos.

Results

The America On Line (AOL) search yielded the following results:

Iowa Consortium for Substance Abuse Research and Evaluation
http://www.uiowa.edu/~vpr/research/organize/subabuse.htm
Coordinates research and knowledge transfer among researchers, assists professionals in the field, and informs public policy makers in the area of substance abuse.
http://iconsortium.subst-abuse.uiowa.edu

Center for Substance Abuse Research (CESAR)
http://www.cesar.umd.edu

Alcohol and Other Substance Abuse Resources
Information, toll-free numbers, and links to resources for alcoholism, dual diagnosis and substance abuse organizations.
http://www.supportpilot.com/substance.html

Substance Abuse Training Project
Contains the MCBAP Approved Calendar of Events, HIV/AIDS Calendar, information on obtaining credentialing and other resources for substance abuse professionals.
http://www.macmhb.org/subabuse/satp.htm

Film Ideas: Smoking and Substance Abuse
Instructional films and videos on tobacco and substance abuse.
http://www.filmideas.com/substance.html

Schizophrenia and Substance Abuse - BibliographyOne page listing of over 80 publications that address schizophrenia and substance abuse.
http://depts.washington.edu/adai/lib/bibs/ps_175.htm

Center On Addiction and Substance Abuse
CASA Survey: Many Dads AWOL in the Battle Against Teen Substance Abuse.
http://www.casacolumbia.org/newsletter1457/newsletter_s

Women and Substance Abuse - A Selected Bibliography
Published online by the Addiction Research Foundation
http://www.arf.org/isd/bib/women.html

Join Together Online: Substance Abuse
Join Together is a national resource center for communities working to reduce substance abuse and gun violence. "Tobacco is the most common substance abuse problem, and the most lethal. Tobacco causes 400,000 deaths each year, killing more people than AIDS, alcohol, drug abuse, car crashes, suicides, and fires combined."
http://www.jointogether.org/sa

Next are the results of the same search using the Yahoo search engine.

National Council on Alcoholism and Drug Dependence, Inc. (NCADD) - advocates prevention, intervention, research and treatment of alcoholism and other drug addictions.
http://www.ncadd.org/

Brown University - Center for Alcohol & Addiction Studies - promotes the identification, prevention and effective treatment of alcohol and other drug use problems in our society through research, publications, education and training.
http://center.butler.brown.edu/

University of Minnesota, Duluth - Center for Addiction Studies - behavioral science research into addictions with emphasis on alcoholism and compulsive gambling.
http://www.d.umn.edu/~jlaunde3/addiction.html

Association for Medical Education and Research in Substance Abuse - multidisciplinary specialists in the field of substance abuse dedicated to improving education about alcohol, tobacco, and other drugs.
http://center.butler.brown.edu/AMERSA

Addiction Research Foundation - develops and makes available knowledge and programs that can be used to reduce the problems caused by the abuse of alcohol and other drugs.
http://www.arf.org/

University of California, San Francisco Drug Dependence Research Center - includes general info, public resources, researchers at DDRC and links to related sites.
http://itsa.ucsf.edu/~ddrc/

National Center on Addiction and Substance Abuse (CASA) - provides access to information, research and commentary on tobacco, alcohol and drug abuse issues, including prevention, treatment and cost data.
http://www.casacolumbia.org

Stop Drugs - information on methamphetamine, links, research, prevention, treatment and enforcement information.
http://www.stopdrugs.org

Project Craft - federally-funded research program for family members and others who are concerned about drug use problems in a loved one.
http://www.unm.edu/~craft

Research Institute on Addictions - a non-profit institute dedicated to research into alcoholism and substance abuse.
http://www.ria.org

Now, let's try Lycos and compare its results with those of AOL and Yahoo:

Center On Addiction and Substance Abuse - A unique think/action tank that engages all disciplines to study every form of substance abuse as it affects our society.
http://www.casacolumbia.org/

The Heart Disease and Cardiology Home Page - The starting place for exploring information on heart disease and cardiology.
http://heartdisease.miningco.com/msubanger.htm

CESAR's Home Page - Frame Version - Issue 47 - Genetic Influences Contribute Slightly More than the Environment to Risk for Alcohol and Nicotine Dependence Issue 46 - Top Television Shows Fail to Portray Negative
http://www.cesar.umd.edu/

HabitSmart Home Page - This Web site has been constructed to provide an abundance of information about addictive behavior: theories of habit endurance and habit change as well as tips for effectively
http://www.cts.com/crash/habtsmrt/

Welcome to ALKY.COM: Your Ultimate Starting Point for Alcoholism, Substance Abuse, Health and More! Looking for Alcoholism?
http://alky.com/

Substance Abuse Research Center - Substance Abuse Research Center funded by This page is funded by National Institute on Drug Abuse (NIDA) and supported at UCLA Department of Psychology Click at the name on the right to see their vita
http://nidalab.psych.ucla.edu/index.html

Center On Addiction and Substance Abuse - A child who reaches age 21 without smoking, abusing alcohol or using drugs is virtually certain never to do so." - Joseph A. Califano, Jr., CASA Chairman and President
http://www.casacolumbia.org/

Substance Abuse Research Program - Call for Proposals: Substance Abuse Policy Research Program 2000 Round VI Contents: Purpose Background The Program Selection criteria Use of grant funds Timetable- $100,00 Timetable-$100,000-$350,000
http://www.rwjf.org/grant/cfp-subab-00.htm

Iowa Consortium for Substance Abuse Research and Evaluation - Iowa Consortium for Substance Abuse Research and Evaluation Projects Affiliates Teaching Treatment Research Evaluation Analysis Links What is The Consortium ? Mission statement How to reach us A list
http://iconsortium.subst-abuse.uiowa.edu/index2.htm

NYS OASAS: The Office of Alcoholism and Substance Abuse Services - New York State Office of Alcoholism and Substance Abuse Services George E. Pataki, Governor Jean Somers Miller, Commissioner About OASAS Message from Commissioner Agency Organization
http://www.oasas.state.ny.us/

Conclusions

Looking through the resultant lists from the searches we can see that while there are several sites listed in common by all three engines, Yahoo, Lycos, and AOL, there are quite a few that only appear on one list or the other. Notice that a web site on heart disease and cardiology appeared in the Lycos search results even though the title and description did not contain the words "substance abuse." Most of the search engines found web sites with the words "substance abuse" either in the title of the web site's URL or in the title or description of the web page. There are several conclusions that can be drawn from these results. First, results of web searches may be as much a result of the web browser employed than the actual content of the pages. Some search engines, in fact, are moving toward listing only those sites that pay a subscription fee, thus effectively eliminating many excellent resources from being listed. The methods of categorizing sites also may preclude their listing depending upon which categories they are listed in. Thus there were no "substance abuse counseling organizations" that appeared in a query of "substance abuse research." A broader keyword search may result in too many irrelevant results but the narrower search terms may not yield all the sites of interest to the researcher.

Secondly, the effective use of hypertext markup language (htm or html) by the web author may have as much effect on page listings as the actual wealth of data and information available on the site. If the web author does not register the site with search engines the best of web sites may go unnoticed. Some search engines "spider" web sites automatically as they are programs written specifically to find web sites, catalog them, and add them to the database of the search engine. Some search engines have built-in "filters" that will not list certain types of pages and this can result in many excellent resources not showing up in web searches.

Thirdly, there are web sites that are not intended to be public access sites. These include sites that contain confidential data sets on clients or patients, proprietary information such as online courses, or intranet sites for universities, agencies, or corporations designed to allow project team members to work together from distant locations. These sites may have the equivalent of an unlisted telephone number: i.e., an unlisted URL that can be accessed only by those who know the Internet address. Some university and other research institutions that operate behind sophisticated firewall security systems may discourage the use of scripting languages that allow their sites to be searched.

Fourth, the number of times the key words or phrases appear in an article may be related to the length of the article or number of pages in the web site, with longer articles or larger web sites containing the keyword or phrase of interest more often than shorter articles or one-page web sites, thus the web site with more pages may achieve higher rankings in some search engines even if a site with fewer pages contains information of equally high quality.

Therefore, results of queries of search engines may be influenced by the following: content of the pages, hypertext markup language authoring skills of the web author, the design of the search engine's web site rating method, the registration of the site with search engines, the selection of keywords used in the search, the number of pages on the web site and number of times the keywords appear, and the intended audience and purpose of the web site. As search engines proliferate, online discussions of their relative merits has become a common topic of discussion on academic lists among such noted professionals in evaluation as Scriven and others (http://www.evaltalk.org). Thus, an effective Internet search may begin with a search of the archives of the listserves of organizations such as the American Educational Research Association, the American Evaluation Association, or journals and organizations for professional librarians and media specialists, such as Online and Online Review. This strategy will more likely maximize the chance of finding web sites with useful, accurate information on the topic of interest.

Suggested Readings

Brem, S. K. & Boyes, A. (2000). Using critical thinking to conduct effective searches of online resources. Practical Assessment, Research & Evaluation, 7(7). Available online: http://ericae.net/pare/getvn.asp?v=7&n=7

Callery, A. (1996). Yahoo cataloging the web. Proceedings of the Conference Sponsored by the Librarians Association of the University of California, Santa Barbara, and Friends of the UCSB Library" (Santa Barbara, CA, April 26,1996). (ERIC Document Reproduction Service No. ED 403886)

Hertzberg, S. and Rudner, L. (1999). The quality of researchers' searches of the ERIC database. Educational Policy Analysis Archives, 7(25). Available online: http://epaa.asu.edu

Schrock, K. (1998). Evaluation of world wide web sites: An annotated bibliography. ERIC Digest. (ERIC Document Reproduction Service No. ED 436187)